Doing space in face-to-face interaction and on interactive multimodal platforms

Jucker, Andreas H ; Hausendorf, Heiko ; Dürscheid, Christa ; Frick, Karina ; Hottiger, Christoph ; Kesselheim, Wolfgang ; Linke, Angelika ; Meyer, Nathalie ; Steger, Antonia

Abstract: In this article, we argue that the spatial environment of everyday interaction has to be understood as a social construct. Co-participants in an interaction make use of the spatial affordances of the interactional architecture around them, and at the same time they interactionally create and maintain spatial configurations. In that sense, they can be argued to be "doing space". Concerning face-to-face interaction, we distinguish between heavily structured material settings that are custom-built for specific types of institutionalized interactions, such as lecture theatres, assembly halls or service encounters; moderately structured settings, such as restaurants, staff rooms or museums; and weakly structured settings, such as public town squares or other settings which provide only minimal assumptions about the interactions that may take place there and their spatial configurations. We extend this analysis to different forms of interaction on interactive multimodal platforms (IMP), where the complexities increase with the different spatial levels of the physical computer screen, the many different spatial levels depicted there, and the increasing difficulties for the interactants to navigate and negotiate the different levels of doing space.

DOI: https://doi.org/10.1016/j.pragma.2018.07.001

Originally published at:
Jucker, Andreas H; Hausendorf, Heiko; Dürscheid, Christa; Frick, Karina; Hottiger, Christoph; Kesselheim, Wolfgang; Linke, Angelika; Meyer, Nathalie; Steger, Antonia (2018). Doing space in face-to-face interaction and on interactive multimodal platforms. Journal of Pragmatics, 134:85-101.
DOI: https://doi.org/10.1016/j.pragma.2018.07.001

Jucker, A.H., et al., Doing space in face-to-face interaction and on interactive multimodal platforms, Journal of Pragmatics (2018), https://doi.org/10.1016/j.pragma.2018.07.001. Pre-publication version. To appear in Journal of Pragmatics 134C (2018) pp. 85-101.

Doing space in face-to-face interaction and on interactive multimodal platforms

Andreas H. Jucker, Heiko Hausendorf, Christa Dürscheid, Karina Frick, Christoph Hottiger, Wolfgang Kesselheim, Angelika Linke, Nathalie Meyer, Antonia Steger

University of Zurich

Abstract

In this article, we argue that the spatial environment of everyday interaction has to be understood as a social construct. Co-participants in an interaction make use of the spatial affordances of the interactional architecture around them, and at the same time they interactionally create and maintain spatial configurations. In that sense, they can be argued to be "doing space". Concerning face-to-face interaction, we distinguish between heavily structured material settings that are custom-built for specific types of institutionalized interactions, such as lecture theatres, assembly halls or service encounters; moderately structured settings, such as restaurants, staff rooms or museums; and weakly structured settings, such as public town squares or other settings which provide only minimal assumptions about the interactions that may take place there and their spatial configurations. We extend this analysis to different forms of interaction on interactive multimodal platforms (IMP), where the complexities increase with the different spatial levels of the physical computer screen, the many different spatial levels depicted there, and the increasing difficulties for the interactants to navigate and negotiate the different levels of doing space.

Keywords: space, face-to-face interaction, 3D virtual worlds, affordances, interactive multimodal platforms (IMP), Second Life, Twitch

1. Introduction[1]

Space is one of the essential contextual dimensions of interaction. Interaction takes place in a spatial context, and it comes into being when the interactants become aware of each other's co-presence. ''Interaction begins when people perceive that they are being perceived'' (Hausendorf, 2012a: 45, with reference to Luhmann, 2005 and Goffman, 1964; our

[1] The present article is the result of the authors' cooperation within the research group "Interactional Space(s)" that comprises different projects concerned with space and spatiality within face-to-face interaction as well as various forms of computer mediated communication, and in particular interactive multimodal platforms. It is part of the University Research Priority Programme (= URPP) "Language and Space" of the University of Zurich that has generously supported our activities (www.spur.uzh.ch). All authors actively contributed both to the research and to the actual writing of this paper. The first two authors took the role of leading authors, all remaining authors are listed in alphabetical order.

translation), and people perceive each other in spatial contexts with a certain distance between each other. The given distance may make it easy to communicate with spoken words or it may restrict interaction to gestures and the like. However, the spatial context is not restricted to the distance between the interlocutors and their body postures. It also includes the multifarious aspects of the physical arrangement of the surroundings. Does the encounter take place indoors or outdoors? Does it take place in a busy street, on a wide, open city square, in a remote country lane, or on top of a mountain after an arduous ascent with the appropriate mountaineering equipment? Or does the interaction take place in a noisy night club, in a doctor's surgery, in a parliamentary assembly hall or at the ticket office in a busy railway station? The spatial configurations in these situations differ considerably, and they have a major impact on the communication that takes place there.

In the built-up surroundings of our modern world, there are many contexts and arrangements that are shaped in such a way as to facilitate interaction or to facilitate certain types of interaction. Ticket offices, for instance, are built in specific ways to enable and facilitate specific forms of interaction, and, in fact, this context raises expectations as to the type of institutionalized and organized interaction that we are likely to encounter there. Restaurants and bars, the living room in a private home or a school staff room are less specific and less institutionalized. They are not purpose-built for certain types of interactions, but in these contexts, too, there are many aspects which enable or facilitate interaction (in order, for instance, to allow for the participants' recreation or some sort of public sociability). Chairs, for instance, are placed at convenient distances and at convenient angles to allow a certain number of people to interact easily with each other. Normally such arrangements are merely taken for granted in linguistic analyses. Or they are taken as fixed entities that make up the context of the interaction.

Here, we advocate a dynamic approach to the spatial context, an approach that focuses on the appropriation of space by the interactants. People "do space" by accommodating space to their communicative needs, by appropriating spatial affordances in specific ways that may or may not have been intended and anticipated by their creators. And people discursively create spaces in interaction. From this point of view, spaces appear to be an interactive and performative achievement rather than a contextual given. However, in today's world, communication is not restricted to the physical world and to the conspicuousness of face-to-face interaction. More and more, we communicate in and through online surroundings and in computer-mediated virtual environments. This raises the question of how space and spatiality play a role in these communicative settings with limited mutual co-presence: How is "space done" in such settings? Is that even possible without a shared physical space?

In the following, we focus on specific interactional arrangements, or architectures-for-interaction as we shall call them (Hausendorf and Schmitt 2016), and the ways in which interactants do space in these social situations. After a discussion of the relevant terminology for our analysis, we will start with social situations in which the setting is heavily structured by the surrounding spatial and architectural affordances so that the interactive achievement of space is done rather inconspicuously, maybe solely in terms of taking the positions offered by the setting. We will then proceed to social situations in which the interactional settings seem to be free from direct spatial affordances and in which the discursive creation of space becomes more and more important and more and more obvious, involving, for instance, language and meta communication. Subsequently, we discuss the role spatial arrangements play in the way users "do space" on interactive multimodal platforms (IMP; Herring 2015). The first example will be taken from *Second Life*, a 3D virtual world popular about a decade ago,

and the second example will be taken from Twitch, in which 3D virtual settings are embedded within the frame of a video-game player broadcasting their gameplay to a wider audience.

2. Terminology

As mentioned above, it has been claimed within the sociology of interaction tradition that interaction depends on the participants' mutual perception of being perceived by each other (see, with further references, Hausendorf 2012a). It is communication under the term of co-presence (as Luhmann 1984 put it following Goffman 1964). This does not require a face-to-face setting in each and every case: Communication over the telephone, for instance, usually allows for both participants to perceive their mutual perception although perception is restricted to what can be heard on the phone. That is the reason why a lack of minimal audible feedback ("hm") typically triggers questions of co-presence ("Are you still there?"). An exchange of messages through WhatsApp, in contrast, could not be seen as interaction in this strict sense, not even if both communication partners are online at the same time, because they do not mutually perceive each other.[2] In this paper, we will leave it open whether this traditional concept of interaction makes sense or whether it proves to be too narrow in the light of computer-mediated communication, as was argued by Dürscheid (2016), who applies the term interaction to situations in which there is a simultaneous presence of the participants in the same (virtual) space and a continuous exchange of turns.

The central notion for our overall argumentation is the concept of "doing x"; in our case "doing space". The "doing" concept is used in different ways. Developed by Harvey Sacks in the 1970s (cf. Sacks 1984), it came to prominence in the form of "doing gender" through the seminal paper by West and Zimmerman (1987: 125), in which they argue "for a new understanding of gender as a routine accomplishment embedded in everyday interaction" (see also Fenstermaker and West 2002). Gender is no longer seen as a biological fact but as a social construct, which is created and re-created in social interaction. People do not behave as man or as woman because of their birth or their upbringing and socialisation but because of their knowledge of what it means to behave as man or as woman. West and Zimmerman distinguish between "sex" as a biological notion, "sex category" as the social ascription of the relevant sex criteria and "gender", which is created in the process of interaction (see also West and Zimmerman 2009). The idea of looking at social concepts in such a dynamic and performative way has been taken up in many different contexts. Schilt and Westbrook (2009), for instance, talk of "doing difference" and "doing heteronormativity"; Haugh et al (2015) of "doing deference"; Georgakopoulou and Charalambidou (2011) of "doing ageing" and Aronsson (2006) of "doing family".

With the notion of "doing space" we refer to the linguistic approach that human interaction is shaped by space and space comes into being through interaction (cf. Hausendorf 2013: 276).[3] Within this paradigm, it is assumed that interactional space results from the interplay of language use, perception, action and bodily arrangement and that it is created and upheld by interaction itself ("interactional space" in the conversation analysis tradition: for instance, Mondada 2009). This is in line with Gibson's notion (1977: 141) of the "basic affordances of

[2] Of course, they can see that the other person is about to write a message, but they do not see what he/she is writing, and they do not see the person.

[3] See also Löw's (2016: 141) concept of "(An)Ordnung", which combines the two aspects of arrangement as order and the process of arranging or ordering.

the environment." Lecture halls, for instance, provide affordances for how the built space might be used for interaction (see above). The same concept has also been applied to virtual space (see Pentzold et al. 2013) in spite of some obvious differences. In a virtual world inhabited by avatars, for instance, objects such as seats are not necessary to provide a comfortable setting for interaction but they provide a "usability cue" (in the sense of Hausendorf 2012b) for the interactants, who imitate physical interaction in virtual environments. Our approach is based on the assumption that these affordances play an important role in the physical as well as in the virtual world and that human interaction is shaped by space and space only comes into being through interaction.

3. Dimensions of spatial settings

In this section, we will present physical-world scenarios to illustrate how we conceptualise the performative doing of space. We start with heavily structured settings, i.e. contexts that are purpose-built for one specific type of institutionalized interaction, such as lecture theatres, consultation rooms and ticket offices. In a second step we look at moderately structured settings in which communication is not the major focus of the physical arrangement but in which communication is frequent and is at least partly structured by the arrangement of furniture or other spatial configurations. Examples for such settings are living rooms or museums. Here, the interactants have to appropriate the given affordances for their communicative goals. Finally, we present weakly structured settings that are largely devoid of pre-structured communicative affordances, such as open city squares or other outdoor locations. In such settings, doing space reaches a different level. It is no longer the physical setting of the surroundings that suggests certain spatial arrangements of speakers and addressees in their interactions. Instead, the participants have to do space in a much more fundamental way by positioning themselves in relation to each other. In all these contexts, we explore the interaction between the interactional architecture, i.e. the affordances provided by the surroundings, and the ways in which interactants appropriate them.

3.1 Heavily structured settings

Heavily structured material settings have in common that they usually have very clear boundaries and that people have very clear expectations as to the type of interaction that takes place there. Lecture theatres, assembly halls and churches are closed-off rooms that are entered through a door, and all the people that are inside typically participate in the communicative activity for which this room has been purpose-built, either as active speakers or as listeners. Even ticket counters are generally marked off as a space reserved for the direct participants in an interaction, e.g. by lines on the floor or by railings or barrier tapes that keep other customers waiting in line at a distance that should prevent them from participating (as listeners) in the interaction between the service provider and the current customer (see below). People who enter these spaces usually know what to expect. In a lecture theatre, they expect academic presentations, lessons or – as the name suggests – lectures. In an assembly hall, they expect political debates, town-hall meetings and the like. In churches, they expect sermons. And at the service counter, they expect an exchange of goods and services. Needless to say that in all such settings other activities can and do occur. But in an obvious sense, they are purpose-built for specific communicative activities, and people share an understanding of what they are. As a result, all these settings are shaped by their interactional

architecture, i.e. their own specific affordances that enable and facilitate the communicative activities for which they were designed.
In the following, we take the interactional architecture of a ticket counter in the counter hall of the main railway station in Zurich as an example to show how "space is done" in such a heavily structured spatial setting.[4] While the complex social space of the station hall allows for a variety of different activities, such as shopping, eating and drinking, meeting or admiring the nineteenth-century architecture of the building, the ticket counters and their immediate surroundings are obviously built for very specific purposes, i.e. purchasing train tickets. They are characterized by a line of counters within a counter hall that is separated from the main hall. The key elements can be seen in Figure 1.

Figure 1: Zurich main station: counter hall with line of customers

The most striking feature of the setting is probably the empty space between the line in front of the counters and the people lined up in some distance from the counters. People are prevented from heading directly towards the counters by a (blue) barrier tape. The setting more or less forces them to line up. It is not before the customers reach the front of the line, and not before a display of vacancy emerges on a monitor, that they are allowed to head for one of the counters. Having entered the counter hall, customers find themselves in a situation in which the social order of waiting and getting their turn is heavily pre-structured by means of architecture-for-interaction, i.e. a spatial design for lining up (which are familiar from airport

[4] The data stem from the project "At the counter" that was initiated by Heiko Hausendorf and Lorenza Mondada (see Hausendorf and Mondada 2017 for more details and for the analysis of openings at the counter).

counters). Participating in an "interaction order" (Goffman 1983) of lining up accordingly becomes a matter of taking already designed and defined spatial positions. As soon as customers take a position at the end of the line behind the barrier tape, they make themselves visible as a waiting client in a most inconspicuous way. Spatial positioning directly turns into social positioning. They are "doing space" by using their bodies as a filler to a slot that has already been provided.

It is such an arrangement that we have in mind when we refer to heavily structured social settings. Such settings allow for a very efficient and economical way of participating in a framework of activities. The barrier tapes can therefore be understood as a physical provision for a social problem, namely that of organizing a turn-by-turn accessibility to the service when the number of possible next clients exceeds the number of actually available counters. The bodily solution in terms of lining up with its social norms of "first come, first served" and "one at a time" (Hausendorf and Mondada 2017: 14) turns into a material arrangement. Thus, the barrier tapes in themselves manifest social norms and expectations in terms of architectural provisions. In this sense, the counter hall setting not only provides affordances for those who enter the counter hall. It is also heavily loaded with social norms as to the kind of interaction order that is expected.

The same holds true for the counter itself. The following piece of data documents in some detail a typical opening of a conversation over the counter (see Figure 2)

Figure 2: Opening of a ticket counter conversation

(1) Ticket clerk: GRÜEzi: WA:S hetet sie gern
'Hello what would you like to have'
Customer: s HALBS, GONtenschwil (.) Retour
'half Gontenschwil [=toponym] return'

In fact, the interaction in (1) starts somewhat before the customer has completed her arrival at the counter (Figure 2). She is greeted by the clerk ("hello") and immediately asked for her request ("what would you like"). There is no greeting pair, and the customer immediately formulates her request ("half Gontenschwil return"). The customer's bodily arrival at the counter coincides with the start of her request ("half").

It is striking how both participants contribute to put across the request so quickly and in such a fine-tuned way that the client's arrival at the counter in fact overlaps with her request. Note

that there is obviously no need for verbal clarifications as to the nature of the interaction or as to the sort of social categories that are relevant for the encounter. It is due to the setting sketched out above that the arriving person can be accounted for as the next client and that the person behind the windowpane can be accounted for as the ticket clerk. So, a lot of highly relevant social expectations have already been put into force when the participants start with opening up the encounter. Bringing into action the spatial and architectural affordances with their bodies, the participants can immediately get to the point.

When interpreting the interactional architecture of the counter in this way, one is led to the assumption that the ticket office manifests an architecture-for-exchange under limited conditions: The architecture of the counter defines the purpose of the encounter as an exchange between tickets (and other goods) and money. In this regard, the counter appears to be the historical precursor of the ticket machine. It reduces the social interaction between the one in front of the counter and the one behind to the purpose of exchanging money for a ticket. It is the expectation of such a kind of reduced sales conversation in terms of exchange that the counter manifests in its material and architectural forms. In this sense it facilitates and speeds up the process of buying a ticket. It is, in the truest sense of the word, a heavily structured setting that allows for innumerable small, efficient and economical interaction episodes between clients and agents. But, contrary to the ticket machine, there are humans involved on both sides of the boundary and there is face-to-face interaction (as focused and directed towards the buying as it may be). Even a heavily structured setting cannot prevent social interaction and its participants from blithely ignoring affordances and expectations and from a sort of “doing space” that goes beyond the taking of already designed places and positions. In most cases, however, it allows for a highly complex type of social interaction by merely doing space in terms of using the material affordances of the setting.

3.2 Moderately structured settings

Under this heading we cover situations in which the communicative affordances are much less specific and allow for a broader range of interactions. In fact, these settings are not purpose-built for specific types of interactions, although there are aspects in the architecture-for-interaction that facilitate communication, and certain types of interactions are more likely than others. Typical examples are restaurants, school staff rooms and living rooms in private homes. These have in common that seats are generally available and arranged in a way that facilitates communication between smaller or larger groups of people. The boundaries are less clear cut than in the railway station setting described above, and multiple communicative events are not only possible but indeed typical for many of these locations. This category also comprises settings in which communication is much less typical but still possible, such as museum exhibitions and supermarkets. In all these settings, communication is possible and it is a regular occurrence, but it is not essential.

We will use a museum exhibition as an example to illustrate how space “is done” in a moderately structured setting. In a certain sense, museum exhibitions are heavily structured, too. They have clear cut boundaries, which almost automatically transform people crossing these boundaries into visitors, guards, museum guides, etc. There is a second aspect of museum exhibitions that makes them appear heavily structured. People have a clear set of ideas about the activities they can expect in the exhibition space. We know how people typically move in a museum – with measured steps, pausing in front of exhibits without moving, etc., and a subset of these expectations refers to communicative activities.

However, and this is an important difference with respect to the ticket counter described above, the interactions which are heavily structured by the architecture and design of the museum space are not interactions among co-present interaction partners, but interactions based on time-persistent signs: exhibits, exhibition texts, and other non-embodied semiotic resources such as graphics, models, banners, the signage system, etc. (cf. Kesselheim 2017). A large number of different affordances enables and facilitates exactly this kind of communication in the exhibition space. In the following we will illustrate this with the example of a museum showcase as an emblematic piece of furniture of the museum exhibition (see Figure 3).

Figure 3: A showcase in a zoological museum

The affordances of the showcase in Figure 3, which is part of the Autochthonous Birds section of a zoological museum, strongly suggest how this piece of furniture is to be used by the visitors. Its most salient feature is arguably its glass casing. The transparency of this material not only enables the visitors to look through it into the interior of the showcase, it can be seen as a strong signal that this is exactly what the visitors are supposed to do here ("Look, but do not touch!"). The relevance of looking into the showcase is further emphasised by the autonomous lighting in the interior of the showcase. Within the showcase, the affordances of the exhibition design point the visitors' visual attention to the exhibits. The shallow depth of the showcase and the positioning of the exhibits on shelves and plinths, for example, bring the exhibits as closely as possible to the eyes of the visitor. In doing so, they suggest that the visitors study the exhibits in all their visual detail. Furthermore, the labels close to the exhibits identify them as representatives of a class of animals, each exhibit representing a different species.[5] In other words, the labels highlight the sign character of the exhibits. They signify that the exhibits are not to be seen as mere objects (plastic sculptures covered with feathers). In looking at them visitors ought to see typical features of a certain class of animals.
While many of the affordances in the exhibition space suggest that the communicative activity one can expect in this location is visitors using exhibits as material signs which allow them to learn something about animal species, it is difficult to find affordances that can be related to face-to-face communication. In this respect, the museum exhibition is a clear case of a

[5] In addition, the small font size of the explanatory labels in the showcase indicates a zone immediately in front of the showcase to be the ideal standpoint for the visual examination of the exhibits, since this is where visitors have to stand if they want to see the exhibits and read the information about them at the same time.

moderately (or even weakly) structured setting.[6] It is, of course, very common to visit a museum not alone but in a small group, but there is very little, if anything, in the architectural affordances that creates the expectation of face-to-face interaction in this setting. The exhibition space is clearly not organised around a communicative hotspot where a common communicative event is staged, such as the altar or the pulpit in a church room. The layout of the exhibition space, however, certainly does not prevent interaction in a group of visitors. There are numerous smaller areas that can easily accommodate smaller conversational groups, e.g. the free spaces in front of the museum showcases. However, if people want to use this area for their face-to-face communication, they have to "do" their interactional space interactively – by bringing about a free-standing conversational group, an F-formation in the sense of Kendon (1990)[7] with the help of their embodied resources: the positioning of their body, the orientation of their heads, mutual gaze, and talk, and, thereby, temporarily ignore the communicative offerings of the exhibition. In addition, the showcase along with its free space does not help (or oblige) people to establish a certain interaction order, in contrast to the ticket counter, which establishes a strict turn-by-turn order. In front of the showcase there is, in general, no need to take turns, and if turns are taken, they follow the general turn-taking mechanisms. This is exactly what the exhibits in the showcase afford: The fact that there are several exhibits in the showcase and that they are placed with a relatively large distance between each other allows several visitors to study the contents of the showcase at the same time, rendering a turn-taking system largely obsolete.

The following data extract (Figure 4a-d) shows how people "do space" in such a moderately structured setting. It illustrates how the setting helps people to transform their activities into museum activities: The only thing they have to do is to ostensibly signal their orientation towards the affordances in the exhibition space we have sketched above.

	
4a. #(1.0)	4b. JE: DA, ZUUNchönig.# *there, wren*

[6] The situation is, of course, different in the museum café and the museum gift shop. These places are organized like other restaurants or shops. Here we focus on the museum exhibition, where the exhibits are displayed.

[7] According to Kendon (1990: 209), "an F-formation arises whenever two or more people sustain a spatial and orientational relationship in which the space between them is one to which they have equal, direct, and exclusive access. Such a pattern can be seen in the circle of the free-standing conversational group."

	
4c. MA: dä ZUUNchönig; *(---)* *the wren*	4d. MA: CHUMM änni. (.)# *come, Änni*

Figure 4a-d shows a group of three adults (Manfred and Ännie with their grown-up daughter, Jeanette) who have walked through the museum space (Fig. 4a) and have now come to a halt. The three visitors show – to each other and to potential observers – that their coming to a halt is motivated by the institutional frame, that it is a case of stopping in order to watch the exhibits, and not a case of stopping due to fatigue or other non-institutional reasons. Jeanette does not stop somewhere in the middle of the passageway, but close to the glass front of a showcase, at a point made more salient by an audio-station with a kind of telephone receiver (Fig. 4b). There she chooses a position which allows her not only to look at the exhibits in the interior of the showcase, but also to read and press the labelled buttons of the audio-station and to hold its receiver to her ear. With her spatial position she clearly demonstrates that she is appropriating the affordances of the exhibition space. This visible orientation to the exhibits makes it possible for her to use a highly elliptical utterance “DA, ZUUNchönig” (‘there, wren’, Fig. 4b) as an invitation to her co-visitors to discover the wren specimen in the showcase. Ännie and Manfred accept Jeannette’s invitation by repeating ‘the wren’ (Manfred, Fig. 4c) and by taking the receiver of the audio-station (Ännie, Fig. 4d). This is what we mean by “appropriation of the affordances”: People activate the affordances in a spatial setting and make them relevant for their ongoing face-to-face interaction (or they display their indifference towards them, marking them as irrelevant for the understanding of their communication).

3.3 Weakly structured settings

In the urban setting, public squares – at least in a European context – are often purpose-built in such a way that they do not impose unified and/or specific interactive actions but ideally function as places where a broad range of different interactions by a broad range of different people should be enabled. Thus, they are a good example of weakly structured settings. The difference from the spatial settings discussed above is that in public squares interactions are less structured by the built environment and the functional expectations that come with it. Instead, they are more structured by interactional routines *in situ*. One short sequence of an encounter on a public square called Werdmühleplatz in Zurich will illustrate this. The example shows the considerable efforts of a walking person to start talking to a sitting person.

The built environment of this square, as far as the stills from the video recording in Figure 5 show, consists of an empty space and two benches set up around trees as well as another bench, which is placed a few meters away. The benches imply possibilities to sit and linger. But apart from that, no interaction is implied by the setting: There is no clear path on which to

walk, there are no functional reasons for stopping or talking; on the contrary, the benches are built in such a way that they encourage bodily positions in which it is difficult to establish an F-formation (Kendon 1990), and above all, there is enough space to allow people to pass each other at a rather large distance. The square is materially structured in a way that there are no probabilities to accomplish something other than just being there.
The sequence of video stills in Figure 5 below shows one person sitting on a bench, focused on a smartphone. Another person is entering the square and walks slowly in the direction of the sitting person. A few steps away, he turns his trajectory towards the sitting person and is seen to talk to her for a few seconds, asking for a lighter.

	
5a: walking person (P_{walk}) shows self-touch, orientation towards sitting person (P_{sit})	5b: P_{walk} shoulder shrug, orientation slightly past P_{sit}
	
5c: P_{walk} steering trajectory towards P_{sit}, orientation slightly past P_{sit}	5d: P_{walk} trajectory and orientation towards P_{sit}
	
5e: P_{sit} starts looking up	5f: P_{walk} stops, focused interaction established

Figure 5: Video stills of an interaction recorded at Werdmühleplatz in Zurich

For the most part of the sequence the sitting person is focused on using a smartphone (5a–5d), not displaying much attention towards the environment. She therefore displays a low

accessibility for face-to-face interaction. The walking person organizes the beginning of their co-presence in an unobtrusive way: His trajectory is one of a typical passer-by, he walks in a straight line between the benches without slowing down or changing direction. While thus displaying a stable trajectory, the movements of his head and arms show more variation. In 5a, he does a self-touch while orienting his head (and thus visual field) towards the sitting person. Directly afterwards (5b), he shrugs his shoulder while turning his head even further to his right side, thus turning his orientation slightly next to the sitting person's left side. This behaviour can be understood as a preparation of his approach when seen in the sequence of the things that follow: Right after his shoulder shrug, he starts turning his trajectory towards the sitting person (5c). It is noticeable that at this moment, he turns his head (and thus his visual field) towards his left side, away from the sitting person to her right side. Only when he has arrived in a frontal position (5d) he stabilizes his visual orientation towards the sitting person who starts looking up soon afterwards (5e). The focused interaction is established at the moment the man stops and the sitting person stabilizes her gaze upwards towards the now standing person (5f).

The walking person displays a structural problem: In order to gather information about the other person (i.e. does she have a lighter?) he has to look at her. But his looking does not seem to be unproblematic. He uses a lot of interactional resources to render his looking somewhat random and not very direct. His self-touch and shoulder shrug seem to emphasize this effort. Only in close spatial proximity does he fix his orientation visibly towards the sitting person and thus starts initiating the focused interaction. As soon as the sitting person looks up directly towards the now standing person, she confirms the approaching person as an interactive counterpart and the focused interaction is mutually accomplished.

This process is not displayed as a clear-cut, unproblematic approach. The interactive challenge lies in the two different ways in which the spatial-social setting is only weakly structured; in the material and in the social dimension. As argued above, the built environment does not make interaction necessary or even expectable. What this example also highlights is the notion of anonymous relations in public settings. These two persons are neither acquaintances nor can they rely on a clear-cut social-functional role that is made expectable through a specific kind of spatial setting as in the examples discussed above (which would be clerk and customer or visitor to a museum). Apart from guessing social categories such as their gender, age and approximate social status, they are unknown to each other and cannot refer to an institutionalized interactional setting for their encounter.

What we see in this example is a considerable effort of one person approaching another person even though his request is minor (see De Stefani and Mondada 2014). Opening up a focused interaction in a materially weakly structured setting puts some extra strain on the interactants. With that in mind, it is no surprise that the data shows that strangers in public squares generally do not talk to each other. In other words, in public spaces, focused interaction is not the most common way of entering social relations. However, there are many interactive instances beyond the focused interaction. The term "unfocused interactions" by Goffman (1963) can help to conceptualize this broad range of interactive processes that happen without creating a common focus. From this perspective, not getting into contact with each other does not imply an absence of interaction or a lack of sociability. On the contrary, a specific feature of weakly structured public squares (and other public spaces, such as streets) is the interactive avoidance of a focused interaction with many other co-present people. This shows how the materially weakly structured setting of an urban square requires the interactants to put an interactive effort into establishing focused interactions as well as putting

considerable effort into making themselves accountable for the contrary, the avoidance of focused interaction. On public squares, interactants can delegate less of the interactive work to their context than in more materially structured settings but have to do more of it themselves. Public spaces can thus be characterized by imposing more strain on the interactants themselves when it comes to organizing interaction processes and at the same time affording a larger freedom of scope in dealing with being in the open.

4. Spatial settings on interactive multimodal platforms

In this section, we extend our analysis to different forms of interaction on interactive multimodal platforms (see Herring 2015; Herring and Demarest 2017). As we will see, the situation here is even more complex than in physical-world settings, given the fact that the internet offers a great number of differently shaped spaces in which diverse possibilities for communication are available: One can communicate via different channels, with varying participants (one-to-one, one-to-many, many-to-many), in different temporal dimensions (synchronous, quasi-synchronous, asynchronous) and by means of various semiotic systems (oral, written, pictorial). And, similarly to the above examined face-to-face interactions in physical space, this kind of communication may be analysed using the methods of interactional linguistics (in the broader sense of "interactional", cf. Dürscheid 2016). Thus, on the internet, too, it is possible that the interaction partners mutually perceive each other, as, for instance, in a web-based video interaction via Skype (cf. Sindoni 2012, 2013) or Google Hangouts (cf. Rosenbaun et al. 2016a). In a text chat, of course, this is not possible in the same way, but the participants refer to each other's utterances as well, even if there are temporal gaps during this turn-by-turn exchange. These gaps may also occur in face-to-face communication; they are very typical for situations such as driving a car, working together on a project, watching TV together, playing video games in a group etc. (cf. Baldauf and Klemm 1997; Ayaß 2012; Piirainen-Marsh 2012). But, compared to this, the spatial settings for online communication are very different, and interactive multimodal platforms (IMPs) provide an even more complex spatial context. IMPs are internet websites that "allow social media users to comment on multimodal content via multiple channels on a single website, and even within a single thread or conversation" (Herring 2015: 398). Complexities vary from one IMP to the next. They minimally involve text chat and one additional modality (such as pictures) but often they involve more than just two. Several levels are to be distinguished here: a) computer users sit in front of their devices in their own physical spatial context (such as a teenager's room, a university computer lab or a cybercafé); b) the computer screen as a technical device provides a second spatial frame of reference; c) the depicted virtual environment provides a third frame of (spatial) reference. We illustrate such IMPs and their complexities with brief case studies of *Second Life* and *Twitch*. These platforms cannot in themselves be classified as heavily or weakly structured spatial settings, but – as we will show – they are host to different settings, or layers of settings, that are in their own ways heavily or weakly structured.

4.1 Second Life

Second Life is an online virtual world in which virtual embodiments, so-called avatars, interact with each other. These avatars can take anthropomorphic, zoological or fantasy shapes or indeed shapes of everyday objects. They navigate through a three-dimensional world of buildings and objects created by the users.[8] In contrast to many other Massively Multiplayer Online Role Playing Games (MMORPGs), there are no tasks or objectives for the residents, who are free to pursue their own interests, to interact with other residents, organise events such as parties, lectures, discussion circles and so on (see Boellstorff 2008; Abdullah 2015; Brookey and Cannon 2009; Locher et al. 2015; Berger et al. 2016). The case study presented here is taken from an exploratory study carried out in 2009 and 2010 and presented in some detail in Berger et al. (2016). It consists of focused *Second Life* events that were strongly framed, with scheduled beginning and ending times, and moderated by a resident who had advertised the event and who generally, at the end of the event, asked for donations in the virtual currency of *Second Life*. The example presented here consists of a lecture in which one resident, let us call him Primo Maximilian,[9] provided instructions on scripting virtual objects in *Second Life*. Figure 6 shows an anonymized and somewhat simplified artist's impression of the scene rather than an actual screen shot. The avatars are depicted in uniform standard shapes in spite of their highly individual and in some cases very elaborate shapes (see also Frohwein et al. 2008). The names of the avatars that normally appear above them have been omitted in order to preserve their anonymity and to declutter the picture. The texts on the screens have been simplified or rendered as generic text.

[8] Following the terminology of the developers of *Second Life*, we distinguish between the physical users who sit at their computers, the residents, i.e. their virtual identities in *Second Life*, and the avatars, i.e. their visual embodiments within the virtual world.

[9] Residents assume names that often project a particular gender. The resident that we used for our participant observation investigation was called Debbie Cyberschreiber and thus projected a female gender. Other names have been omitted altogether or changed, but we have preserved their projected gender. Debbie had a notecard on her declaring her research interest. Nevertheless, there are some ethical challenges with this data because we do not have the necessary permission from all the participants. Therefore, we refrain from providing a screen shot but present an artist's impression of the scene that we discuss here (for details see Berger et al. 2016).

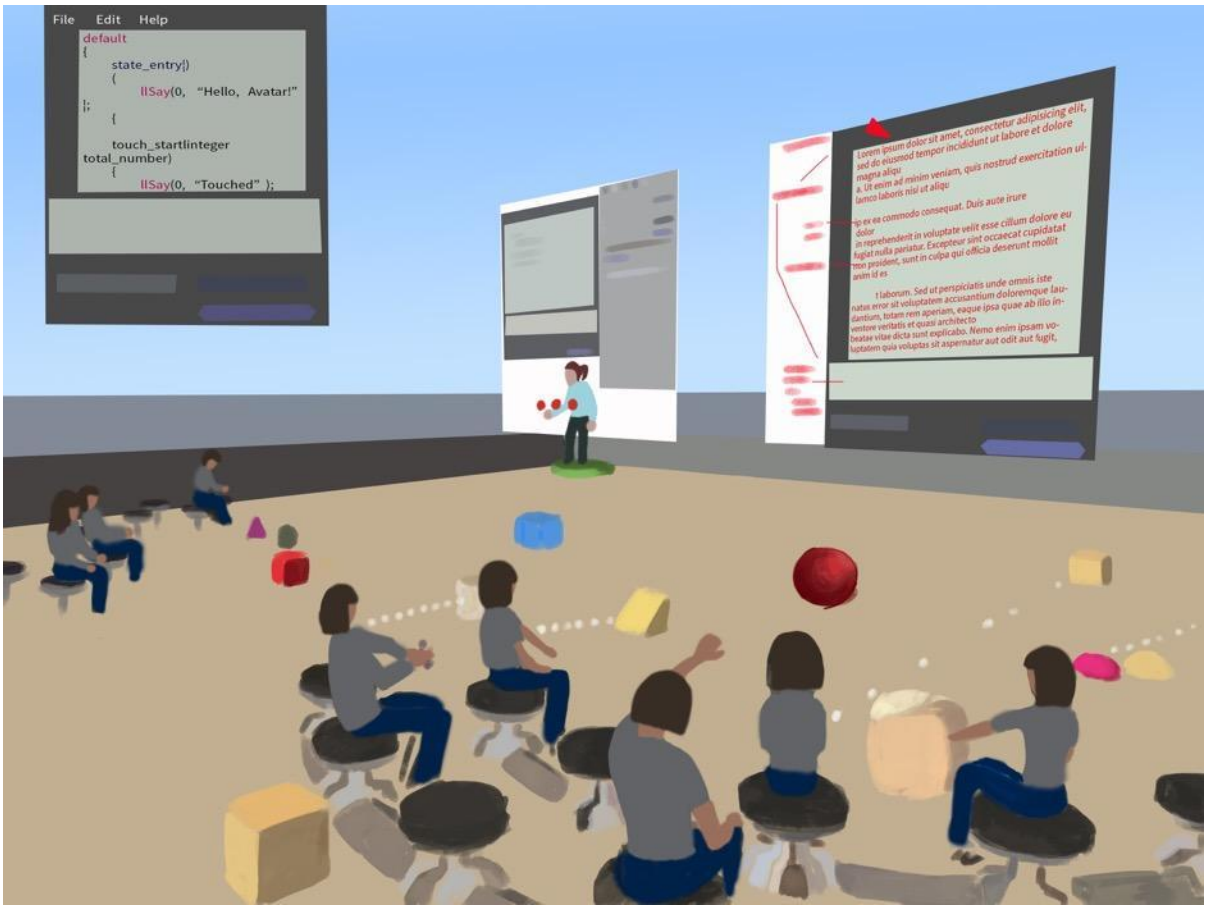

Figure 6: Lecture theatre in *Second Life* (artist's impression with anonymised and standardised avatars, Berger et al. 2016: 90; ©2015 Katrin Jucker)

The lecture theatre depicted in Figure 6 has some similarities with a lecture theatre in physical life. There are a number of seats arranged in a semi-circle around a podium on which we can see the lecturer, Primo Maximilian. Behind Primo's back there are two screens with instructions and sample scripts. The avatars in the audience have assumed a sitting position, while Primo's avatar is standing upright and faces the audience. In physical life, such positionings are comfortable for the audience and they ensure that all the members of the audience can see and hear the lecturer. In the virtual reality of *Second Life*, the position of the avatar is largely irrelevant for the actual purpose of the user learning something about the scripting, but the residents are clearly "doing space". They imitate the spatial positionings of physical-life lectures in order to evoke the social frame of lecturing.

The communicative patterns in this situation are highly complex, and in fact it was not easy for the user navigating our participant-observer resident, Debbie Cyberschreiber, to keep track of all the different channels. The lecturer used a chat window (not visible at the moment captured by Figure 6) to give instructions. At the same time sample scripts appeared on the screen behind Primo, and additional information and instructions were displayed on the screen on his left side. Debbie's user had to write script into the window on the top left of the screen. This window is part of the user's computer screen and not part of the lecture theatre. The residents are trying to activate the prisms in front of them. The dotted lines visualise the interaction between each resident and his or her object.

The interactional architecture created for this virtual lecture theatre makes a clear distinction between spatial affordances that appear to be essential to evoke the communicative event of a lecture, i.e. the relative positioning of the participants, the position of the screens, the seating and standing posture of the audience and the lecturer respectively on the one hand, and the incidental affordances that facilitate the event in physical-life lecture theatres on the other, such as walls and a roof to keep out the glaring sun or inclement weather and the noise from nearby streets, etc. Moreover, the distance between the lecturer and the audience appears to be more than would be easily manageable without audio amplification in physical life. This lecture theatre, therefore, highlights the dual function of spatial affordances. On the one hand, they facilitate and enable a certain type of interaction, here a lecture, and on the other they flag what is going on in this context as exactly this type of interaction. In the virtual reality of *Second Life*, the facilitating aspect is backgrounded. There is no physical need for avatars to sit down or to face the lecturer. And at the same time the flagging function is highlighted. By doing space in the appropriate way, the residents perform the activity of a lecture. So, strictly speaking, doing space in this case has two meanings: On the one hand, as described above (section 2), this concept means that space comes into being through interaction (see the text chat in Figure 7). On the other hand, *Second Life* residents imitate physical space (in our case a lecture hall, and therefore a heavily structured setting). This becomes apparent in some of their activities (such as the fact that they position their avatars in a way that they face the speaker). But the residents do not only move in the architecture provided by the depicted scene and created by their activities as well, they also create interactional spaces. Thus, doing space in this context means doing interactional space and doing physical space.[10]
Figure 7 gives an example. It is a short extract from Primo's lecture. The first four and the last of his "utterances" are only preceded by "Primo". These appear to be pre-fabricated units that Primo's user can post quickly and easily during the lecture, while the remaining three are preceded by "Primo Maximilian". These appear to be utterances that this user actually types in real time during the lecture. Such utterances often respond to what members of the audience do or say.

[7:44]	Primo: First we need to make a place for our scripts to live.
[7:44]	Primo: Look at this slide here on the left.
[7:44]	Primo: It shows the basic window that opens when you make or modify a prim.
[7:44]	Primo: I refer to it as the Prim Editor Window, or sometimes as the Prim Builder Window.
[7:45]	Primo Maximilian: notice which one of the top 5 buttons is highlighted
[7:45]	Primo Maximilian: that the editing widget button
[7:46]	Primo Maximilian: the other one that gets used a lot is the creation widget button which is the 4th from the left button
[7:46]	Primo: If you hear me say that, this is what I'm talking about.

Figure 7: Extract from chat window (pseudonym speaker indication; scripting class – basics, recorded on November 20, 2009)

[10] Note that the term "resident" also refers to physical space, thus it also belongs to the concept of doing (physical) space.

Primo's references to "this slide here on the left", "one of the top 5 buttons" or "the 4th from the left button" create spatial orientations for the audience, which has to figure out the proper point of reference for what is left or top because they navigate not only within the lecture theatre but also on their computer screens.
Thus, the virtual lecture hall depicted in Figure 6 inherits many of the features of the interactional architecture of a heavily structured setting in the physical world. It adopts these features to create communicative spaces that are familiar to the participants, but it does so selectively. It picks out those features that are salient not so much as affordances but as flags or signals which provide the communicative frame for the interaction that takes place in this locality.

4.2 Twitch

Twitch is a livestreaming platform and as such an even more complex IMP than *Second Life* regarding the communicative setting and the spatial layering. Whereas *Second Life* is an IMP in its own right, on *Twitch*, 3D virtual worlds are embedded in a streaming window, which is part of the layout of the platform itself. Through this streaming window, *Twitch* allows video game players, called streamers (cf. Twitch 2017), to broadcast their game play to a potentially global audience in real time. Furthermore, in this particular kind of live streaming environment, all participants – whether it be streamers, co-players or viewers – can communicate with each other through various channels in which they can make use of different modalities (e.g. speech, writing, images, gestures, gaze), depending on the affordances of the game, the streaming platform and various additional third-party tools.
Figure 8 shows a screenshot of the *Twitch* interface as users see it while watching a stream. On the left side, we see a menu bar (1), where users can navigate through channels, access their *Twitch* profile, see their messages from other platform users etc. To the right of that, we see the title of the stream on top (2), the streaming window in the middle (3) with several buttons and statistical information right underneath it (4), as well as the topmost part of the channel description (5). To the right of the streaming window, we find the platform chat (6), in which all logged in participants can actively engage in communication via writing and emojis. In the streaming window, apart from the depicted 3D virtual world of the game (7), in which we can sometimes also see other players' avatars or so-called NPCs (non-player characters), we can also see the in-game chat window (8), which only pops up for a few moments when co-players (e.g. team members or enemy players) engage in communication or when system messages are generated. Right underneath the in-game chat we see a pop-up of a follower alert (9). This, too, is only visible for a few moments before it disappears again. In the top left corner of the streaming window, we find another example of such an overlay (10), which, here, shows the name of the song that is played in the background of the stream. In this case, however, it is not a pop-up, but is displayed throughout the whole stream. Finally, in the lower right corner of the streaming window, we can see the streamer sitting in a chair in their actual physical environment in front of their computer (11).
This overlay is created by using a webcam image with a green-screen behind the streamer to block out unwanted depictions of the actual physical space he/she is located in. Nevertheless, as a viewer, one gets to see the streamer's upper half of the upper body, as well as their face and at times gestures if the hands come into the camera's range. However, not all streamers use the webcam feed and some decide on just broadcasting the game without showing

themselves to their audience. The same goes for all kinds of third-party pop-ups and overlays which are used by a majority of streamers, but not by all.

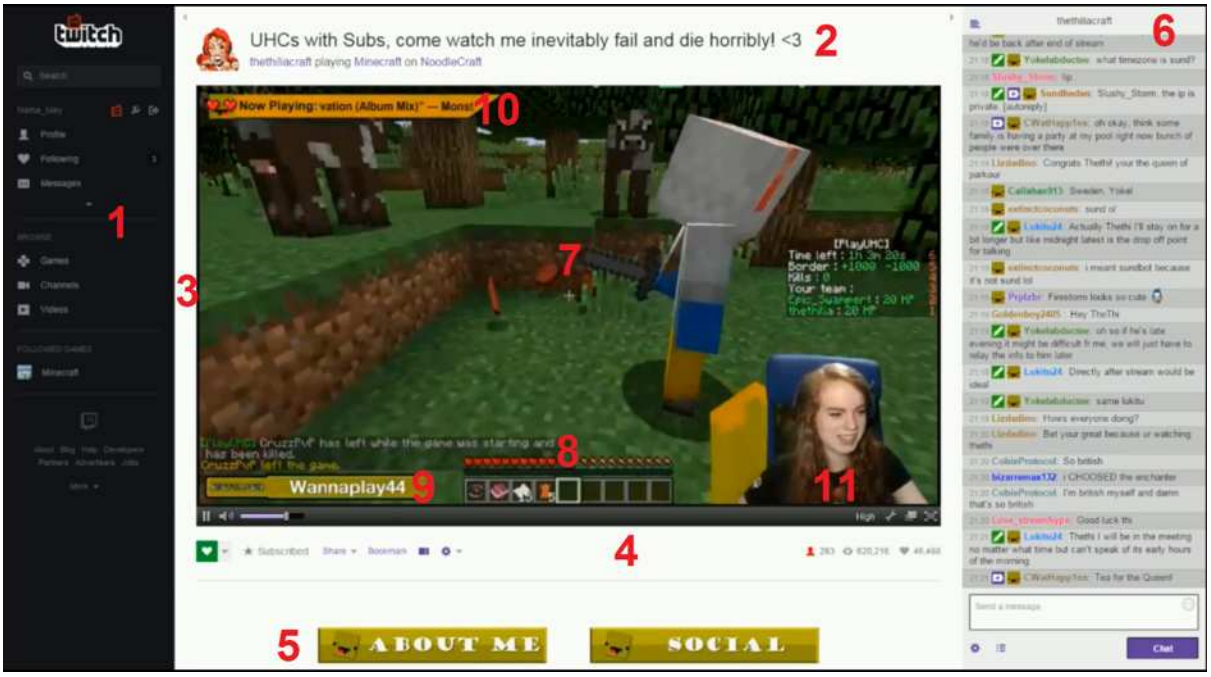

Figure 8: The *Twitch* interface from a viewer's perspective.

Thus, in addition to the physical spatial context of the computer user as a first spatial frame of reference and the computer screen as a second one (see section 4), there are various sub-layers that go above and beyond the third spatial frame of the depicted 3D virtual world. From a viewer's perspective, what can be seen on the screen is not only the depiction of a virtual environment originating in one source, but a combination of different spaces from different sources (physical space of the streamer, virtual game world, graphical layovers from third party services etc.). These spatial layers were then merged and arranged by a broadcasting software and represented as one in their output, the content of the streaming window, which itself is placed in a given slot within the layout of the streaming platform's channel page.
All of these layers can, to some extent, be navigated by streamers and other participants alike. A streamer, therefore, is not only acting as the host of the broadcast and the player of the game, but he/she is doing space by acting also as the navigator between these spatial layers, always verbally and non-verbally indicating which spatial frame of reference is relevant at a specific point in time, which also helps viewers follow the flow of the stream as a whole.
Furthermore, since the end product that a viewer sees on *Twitch* is essentially a composition of the platform and the content of the streaming window, it is impossible to characterize this kind of platform as either weakly, moderately or heavily structured. Rather, a distinction has to be made between the platform *Twitch* on the one hand and the embedded streaming window with its layered content on the other hand. The content on *Twitch* that is placed around the streaming window is to a great extend prestructured: The chatroom is a given and viewers expect to communicate via this channel with each other and the streamer through writing and static images, so-called emojis. The information area is also prestructured regarding its layout. The individual elements are ordered in three columns, and their length differs according to the amount of content. Even though the actual content is created by the streamer, the information

area on most streaming channels tells viewers about the streaming schedule, the computer setup, the social media presence of the streamer, the chat rules and how to donate money. Furthermore, the title of the stream, as well as other elements, such as the viewer counter and various buttons, are also placed in a specific area, which cannot be changed by streamers.
With reference to the notion of heavily, moderately and weakly structured settings in physical space, the IMP *Twitch* as such can be said to have a heavily structured layout. However, the streaming window, even though placed in a specific spot in this layout, is a different matter. Streamers can freely choose what they want to stream and how they arrange different layers such as the game and its 3D world, the webcam, and third-party pop-ups and frames. Thus, in a video game stream, even though the depicted content of the 3D world itself may be weakly or heavily structured, the layered content of the streaming window is considered weakly structured, as streamers can freely choose to integrate a variety of tools and communication channels.
The case study presented here is from a *Twitch* stream of two teams playing against each other (the so-called gameplay) in the 3D virtual world of *Minecraft*. In this particular game mode, two players per team have one hour to gather as many valuable resources as they can in order to build tools, weapons, and enhancement objects for encounters with enemy teams. Usually, streamers have a side-by-side multiscreen setup, so they can see the game on one screen and the platform chat on another screen, which results in streamers looking straight ahead when focusing on the gameplay, but shifting their head position and gaze to the side when they read platform chat messages.
Example (2) shows an example of a streamer-to-co-player and streamer-to-viewer interaction and provides a transcript of spoken utterances by the streamer thethiliacraft. The example illustrates the situation of shifting between the two screens and essentially between two spaces: the 3D virtual environment of the game and the chatroom on the streaming platform. In lines 1-4, she explains to her teammate what next steps they should perform in the game, namely chop trees to find apples before moving the avatars close to the centre point of the 3D virtual world to dig into the ground there (line 2, “zero zero”). Her focus is on the gameplay screen and thus the ludic component of her stream at that moment, because she is talking to her co-player about gameplay strategy while also repositioning her avatar to chop down a tree, so she needs to see what she is doing with the virtual character. However, in line 4, she begins to shift her gaze away from the virtual world and looks at the secondary screen to her left to focus on the chat and thus the social component of her livestream. Her avatar is still chopping down a tree, but since the angle of the avatar does not have to be repositioned for a few seconds, she has time to shift her focus to the chatroom and react to what the chat participants wrote. First, she provides a positive answer to a question posed by user nick, who wanted to know if he has to subscribe to the channel in order to gain access to the server for the multiplayer game. Second, she reads aloud a message written by another user explaining what the term red shirt refers to. The red shirt discussion has been going on for a while at this point in the broadcast.

(2) Streamer switching spatial layers (transcript of spoken discourse).

1	thethiliacraft:	so chop down as many trees as we can [looks into camera]
2		um in a few minutes we'll head as close to zero zero as
3		we can and we're just going to get a bunch of apples and
4		then we're going to start digging [head turned to the left]

5	uh yes nick
6	red shirts is a star trek reference

Thethiliacraft is doing space here on several levels, with her avatar, her body, and interactionally. Similarly to the example of *Second Life*, in the 3D virtual world of *Minecraft*, the physical world with objects like trees and apples is imitated and the avatar has to be positioned in a specific manner in order to chop down such a tree, which is something that a human being in the physical world would have to do as well. Additionally, thethiliacraft is also doing space when she consciously shifts to different spatial layers outside of the 3D virtual world, or in other words when she shifts from one virtual space to another virtual space. On the one hand, she shifts her gaze and by that indicates that her focus shifts away from the depicted virtual world of the game to the platform chat. Experienced viewers will notice this gaze shift and will likely interpret it as a switch to another spatial layer, expecting the streamer to comment on something that has been going on in the chatroom. On the other hand, through speech, she first refers to different entities and what to do with those in the virtual world (i.e. chopping trees, move to certain coordinates, collect apples, dig into the ground) before she refers to something completely different (i.e. nick, red shirt, and star trek). Since her teammate is neither called nick, nor has he asked anything or said anything about red shirts, it becomes evident that thethiliacraft has temporarily shifted to another layer, leaving the ludic virtual space of the game and focusing on what is going on on her second screen, i.e. the social virtual space of the chatroom.

This so-called cross-modal interaction[11] (cf. Rosenbaun et al. 2016a) is typical of *Twitch* broadcasts where a streamer primarily engages in communication via the mode of speech, whereas viewers are restricted to the written mode. Cross-modal exchanges are a relative novelty and present an extension of the mode-switching model applied by Sindoni (2012, 2014), who looked at video-mediated communication via *Camfrog*.[12] However, while mode-switching refers to users alternating between spoken and written discourse throughout a video chat event, Twitch participants are not switching between modes, as a streamer usually communicates only via speech and viewers engage in written chat communication without the possibility to even switch between spoken and written modes. Thus, by chat participants asking questions in writing and streamers answering via speech, they engage in cross-modal communication. In the same manner, the spatial switches are a further extension of the cross-modality concept. Whereas in video chat situations, a participant's focus is anchored in one tool or one virtual space, i.e. the Skype interface, in video game livestreaming such as on *Twitch*, participants focus on different sources or virtual spaces and have to switch between those to communicate effectively, in this example the 3D virtual world and the platform chatroom.

In this example, the spatial layers involved are only virtual, but a streamer can also shift from the ludic virtual space of the game to the local physical space around him/her, for example, to pick up a mug or talk to a person entering this physical space (cf. also Rosenbaun et al. 2016b for a discussion on such interferences from their domestic sphere while people are engaged

[11] Cross-modal interactions are defined by Rosenbaun et al. (2016a: 29) as "interactions in which the production modality is different from the interlocutor's feedback modality, in the same communicative event and in synchronous fashion."

[12] Next to mode-switching, instances of cross-modal interaction could already be found in Sindoni's (2014) data set, e.g. in examples where a user answered via speech instead of writing to a question that was written by another user.

in video chats on Google Hangouts). Space, in this livestream setting, is structured largely by the arrangement of screens in the streamer's physical space, the layout of the platform, the arrangement of sources in the streaming window, and the spatial configurations within the world of the game. Consequently, interaction is equally shaped by the participants' access to and the affordances of the physical space of the streamer, the platform (chatroom with written text and emojis), the game, and the third-party software (e.g. Skype, Teamspeak etc.). For example, while audience members who are not part of the player circle are restricted to writing their messages in the chatroom on the platform, a streamer's teammate can also engage in communication via spoken language through the voice chat to which he/she has access as an inner-circle member. This explains why streamers have to shift between the spatial layers of the game and the platform chatroom if they want to interact with their audience, and it shows how space influences interaction in this particular setting.

Thus, similarly to the example from *Second Life* in section 4.1, where users had to figure out different points of reference by others, in a *Twitch* broadcast, participants always have to interpret utterances according to which spatial layer and also which object/event in the virtual world or which message in the chatroom they refer to. These kinds of live streaming broadcasts provide a complex setting where every participant has to be aware of the different spatial layers and how they are navigated and made relevant, i.e. how space is done, in verbal and non-verbal communication.

5. Discussion and conclusion

In this article, "doing space" has been introduced as a phrase to emphasize that the spatial environment of interaction should be understood as a social construct rather than a physical given, while making use of partly physically and partly technologically manifested architectures-for-interaction. The category of space is then in line with other categories used to highlight the constitutive, constructive power of communication. We are not the first scholars who have tried to apply this approach to space and spatiality (see section 2 above). The so-called "spatial turn" in the humanities and social sciences (cf. Döring and Thielmann 2008) is, roughly speaking, the result of this attempt. Nevertheless, one might wonder if there has been a corresponding spatial turn in linguistics, too. Apart from mostly cross-linguistic approaches to spatial cognition and spatiality in natural language(s) (cf. Levinson and Wilkins 2006), it is the conversation analytical approach that has drawn our attention towards the making of space. It has supported the view that spatial parameters of the speech situation, the co-participants' *here*, have to be treated as constructs interactionally achieved through embodied talk (cf. Mondada 2009). This approach comes close to our starting point. But we differ from it in at least two substantial ways as we have shown above: First, we do not restrict ourselves to the domain of face-to-face interaction but include computer-mediated communication on interactive multimodal platforms. In doing so, we ask for a broader concept of doing space that allows to account for both the role of space among co-present participants and the role of space among users connected via a shared screen (see the discussion of keyboard-to-screen communication by Jucker and Dürscheid 2012). Accordingly, we have discussed different means and forms of doing space in two interactive multimodal platforms (IMPs), *Second Life,* and *Twitch*.

Second, we do not treat space as being communicatively created from scratch. Instead, we have argued that what we call architecture-for-interaction is a powerful resource which allows for more or less (pre)structured social settings. We have provided insights from different social

settings with – as far as face-to-face-interaction is concerned – differently structured spatial structures (heavily, moderately and weakly). Additionally, we turned to computer-mediated communication (in a broad sense) in order to demonstrate how spatial parameters come into play in online surroundings and shared virtual environments. As a result, among others, we have shown that communication in virtual settings obviously uses (the graphical suggestion of) architecture-for-interaction as a flag or a recognitional for certain activity types, communicative genres and formats. This illustrates how the indexical power of social settings in face-to-face interaction is used within IMPs.

As for face-to-face interaction, we have provided evidence from conversation over the counter at railway station ticket offices (as an example of a heavily structured social setting), from exchanges in front of showcases in exhibition halls (as an example of a moderately structured setting) and from mostly unfocused social interaction within the public sphere of an open public square (as an example of a weakly structured social setting). Concerning conversation over the counter, we have shown that it usually does not take any verbal efforts for the co-participants to make use of spatial affordances. Doing space is realized by taking a certain spatial position in the counter hall (lining up or walking towards the counter) and thereby signalling a certain social position (namely that of a prospective client). So, place is "done" largely due to the use and enactment of material prerequisites. And it is not by chance that the institutionalization of service talk manifests itself in a heavily structured architecture-for-interaction.

In contrast, the inspection of the museum exhibition has provided evidence that this setting primarily allows for the exploitation of non-embodied semiotic resources of very different kinds (among which are written texts) – a type of communication that does not necessarily imply focused interaction with co-present others but, of course, does not exclude it. Space can be done alone, i.e. without a co-present partner and merely depending on the communicative usability (and readability) of the exhibition, and it can be done in terms of an interactive achievement. Participants who walk through an exhibition together can adjust their interactional space(s) according to the affordances of the exhibition (for instance, grouping themselves in front of a showcase), but they can also produce common spaces of co-orientation, co-ordination and co-operation more or less independently from such affordances (as our analysis has shown). The setting obviously allows for different spaces to be done according to the participants' actual communicative (and non-communicative) needs. But they still act as visitors from the moment they enter the exhibition hall.

Finally, there are settings, such as urban squares, that are largely free from activity type and purpose-built prerequisites. To some extent, it is their social rationale to offer an open space suggesting merely to spend time in it; maybe alone, maybe together with others doing the same, maybe for a short and fleeting period of time, or maybe for a longer stay. Compared to heavily and moderately structured settings, focused interaction is merely an option among other options, including non-focused interaction and all kinds of civil inattention in Goffman's sense. As was illustrated in our case study of a brief encounter between two strangers at Werdmühleplatz in Zurich, doing space in this setting means to share a large space together with others without engaging in face-to-face encounters – it means, for instance, to avoid ending up in an F-formation just by chance.

To sum up, our examples drawn from face-to-face interaction make clear that doing space appears as a result of applying embodied practices within an architecturally differently pre-structured setting. Bodily and architectural resources that are made use of in highly inconspicuous but enormously effective ways seem to be essential for doing space – and

seem to be more relevant than linguistic resources in many cases in which no verbal forms are needed (for instance, local deixis or toponyms) to do space. However, communication is no longer restricted to face-to-face interaction, understood as communication among co-present participants who can perceive their being perceived by others mutually and synchronically (Goffman 1964 and introduction above). The spread of writing and reading ("literacy") over the centuries and the emergence of "secondary orality" (Ong 1982) in radio and television have long since accustomed us to communication beyond the closeness of face-to-face encounters. So, communication has overcome the restrictions of spoken interaction taking place within space-time units of those present. One could surmise that doing space in the sense we have just illustrated has, accordingly, become unnecessary and superfluous under these conditions, since the commonly shared space of present participants has been replaced by a spatial split-up between writer and reader, producer and recipient. However, in the present article, we have claimed instead that it seems more appropriate to assume that doing space remains relevant but takes new forms. For instance, doing space might be conceptualised in terms of referring to space when we think of spatial references to places in letters or postcards; or in terms of describing space when we think of spatial reference in literature. Especially in the case of fixed and stationary texts (such as traffic signs or inscriptions), readers do space in terms of inferring the meaning from the concrete spatial surrounding (Scollon and Scollon 2003).

This picture becomes much more complex when we turn to various forms of interaction on interactive multimodal platforms, as we have argued on the basis of examples from *Second Life* and *Twitch*. Within these virtual environments, a user (in front of his or her computer) communicates with (an)other user(s) (in front of their computer(s)) by sharing the same depicted content on each screen, for instance, to make writing appear not only on the user's own but also on his or her partners' screen by typing on the keyboard. As far as *Second Life* is concerned, this commonly shared content shows a virtual world, and the users communicate with each other through acting within this virtual world. They do so by enacting virtual embodiments of themselves (avatars). It does not come as a surprise that doing space is then a job for those avatars as soon as they start to come into contact. The users create and imitate virtual architectures-for-interaction as flags or recognitionals for the type of interaction they want to get into: for instance, a virtual lecture theatre for the activity type of lecturing. Thus, they do both interactional and physical space.

The same holds for the case of *Twitch*. In a certain sense, the users act as if they were engaged in face-to-face interaction. But that is not the whole story. Apart from the virtual setting which is visible (and able to be animated and worked) on the screen, there is the spatial setting of the user(s) in front of his or her (or their) screen(s) and, as a part of this environment, the screen usually shows much more than what is actually shared with others. Interestingly, the users' physical spaces can also become a part of the interaction. At almost any time, users can refer to what is part of their own screen – as was shown in *Second Life* when users change between different points of reference due to their navigation within the virtual lecture theatre or due to their navigation on their own computer screen. As we have illustrated, the setting is dramatically broadened in the case of *Twitch* where we get an even more complex variety of communicative channels and technical devices all represented on the screen – or on side-by-side multiscreen setups. This complexity requires a lot of space to be done concretely in terms of shifting attention between different screens (and parts of screens) by gaze and head movements. Accordingly, there is a special need for co-orientation, co-ordination and co-

operation among the users in order not to get lost in the large variety of possible spatial orientations (see our discussion in section 4.2).

It was the aim of this article to shed some light on the many ways in which space is done within and by communication and how doing space differs according to different media-specific conditions. We would like to argue that much research regarding multimodality and space still remains to be done (but see Jewitt 2016 or Keating 2016, for instance), in particular with respect to new forms of interaction in virtual 3D settings, which obviously challenge our understanding of computer mediated communication that has long been treated as some sort of electronic variety of written discourse without seriously taking into account the enormous properties of digitalization. Doing space could be a key concept to explore these properties as important and fascinating resources of communication.

References

Abdullah, Ashraf R. (2015). Second Life: Language and virtual identity. In: Alexandra Georgakopoulou and Tereza Spilioti (eds.). The Routledge Handbook of Language and Digital Communication. London: Routledge, 273-288.

Aronsson, Karin (2006) Doing family: An interactive accomplishment. Text & Talk 26(4/5), 619-626.

Ayaß, Ruth (2012) Communicative activities during the television reception. In: Ruth Ayaß and Cornelia Gerhardt (eds.). The Appropriation of Media in Everyday Life. (Pragmatics & Beyond New Series, 224). Amsterdam: John Benjamins, 23–46.

Baldauf, Heike, and Michael Klemm (1997) Häppchenkommunikation: Zur zeitlichen und thematischen Diskontinuität beim fernsehbegleitenden Sprechen. Zeitschrift für Angewandte Linguistik / GAL-Bulletin 2, 41-69.

Berger, Manuel, Andreas H. Jucker and Miriam A. Locher (2016) Interaction and space in the virtual world of Second Life. Journal of Pragmatics 101, 83-100.

Boellstorff, Tom. (2008) Coming of Age in Second Life. An Anthropologist Explores the Virtual Human. Princeton and Oxford: Princeton University Press.

Brookey, Robert Alan, and Kristoph L. Cannon (2009) Sex lives in Second Life. Critical Studies in Media Communication 26.2, 145-164.

De Stefani, Elwys, and Lorenza Mondada (2014) Reorganizing mobile formations: When “guided” participants initiate reorientations in guided tours. Space and Culture, 17/2, 157-175.

Döring, Jörg, and Tristan Thielmann (eds.) (2008) Spatial turn. Das Raumparadigma in den Kultur- und Sozialwissenschaften. Bielefeld: Transcript Verlag.

Dürscheid, Christa (2016) Neue Dialoge – alte Konzepte? Die schriftliche Kommunikation via Smartphone. Zeitschrift für Germanistische Linguistik 44(3), 437-468.

Fenstermaker, Sarah, and Candace West (eds.) (2002) Doing Gender, Doing Difference. Inequality, Power, and Institutional Change. London: Routledge.

Frohwein, Stefan, Christof Goldhammer and Anna Eggers (2008) Sprache und Kommunikation in *Second Life*. Available at: http://www.mediensprache.net/archiv/pubs/4047.pdf.

Georgakopoulou, Alexandra, and Anna Charalambidou (2011) Doing age and ageing: Language, discourse and social interaction. In: Gisle Andersen and Karin Aijmer (eds.).

Pragmatics of Society. (Handbooks of Pragmatics 5). Berlin/New York: De Gruyter Mouton, 31-52.
Gibson, James J. (1977) The theory of affordances. In: Robert Shaw and John Bransford (eds). Perceiving, Acting, and Knowing. Hillsdale: Lawrence Erlbaum, 127-141.
Goffman, Erving (1963) Behavior in Public Places. Notes on the Social Organization of Gatherings. New York: The Free Press.
Goffman, Erving (1964) The neglected situation. In: John J. Gumperz and Dell Hymes (eds.). The Ethnography of Communication. Menasha: American Anthropological Association, 133-136.
Goffman, Erving (1983) The interaction order. American Sociological Review, 48(1), 1-17.
Haugh, Michael, Wei-Lin Melody Chang and Dániel Z. Kádár (2015) "Doing deference": Identities and relational practices in Chinese online discussion boards. Pragmatics 25(1), 73-98.
Hausendorf, Heiko (2012a) Der Hörsaal als Interaktionsraum. Ein exemplarischer Beitrag zur Archäologie der Vorlesung. Bulletin suisse de linguistique appliquée 96, 43-68.
Hausendorf, Heiko (2012b) Über Tische und Bänke. Eine Fallstudie zur interaktiven Aneignung mobiliarer Benutzungshinweise an der Universität. In: Heiko Hausendorf, Lorenza Mondada und Reinhold Schmitt (eds.). Raum als interaktive Ressource. Tübingen: Narr, 139-186.
Hausendorf, Heiko (2013) On the interactive achievement of space – and its possible meanings. In: Peter Auer, Martin Hilpert, Anja Stukenbrock and Benedikt Szmrecsanyi (eds.). Space in Language and Linguistics: Geographical, Interactional, and Cognitive Perspectives. Berlin/Boston: de Gruyter, 276-303.
Hausendorf, Heiko, and Lorenza Mondada (2017) Becoming the current client. A study of openings at Swiss railway station counters. Arbeitspapiere des UFSP Sprache und Raum (SpuR). SpuR Nr. 05, March 2017. Zurich. Available at: http://www.spur.uzh.ch/de/research/publications.html.
Hausendorf, Heiko, and Reinhold Schmitt (2016). Interaktionsarchitektur und Sozialtopographie: Basiskonzepte einer interaktionsanalytischen Raumanalyse. In: Heiko Hausendorf, Reinhold Schmitt und Wolfgang Kesselheim (eds.). Interaktionsarchitektur, Sozialtopographie und Interaktionsraum. Tübingen: Narr, 27-54
Herring, Susan C. (2015) New frontiers in interactive multimodal communication. In: Alexandra Georgakopoulou and Tereza Spilioti (eds.). The Routledge Handbook of Language and Digital Communication. London: Routledge, 398-402.
Herring, Susan C., and Bradford Demarest (2017) "I'm the first video Voicethread--it's pretty sweet, I'm pumped": Gender and self-expression on an interactive multimodal platform. Alsic 20.1.
Jewitt, Carey (2016) Multimodal analysis. In: Alexandra Georgakopoulou and Tereza Spilioti (eds.). The Routledge Handbook of Language and Digital Communication. London: Routledge, 69-84.
Jucker, Andreas H., and Christa Dürscheid (2012) The linguistics of keyboard-to-screen communication: A new terminological framework. Linguistik online 56, 6/12. Available at: http://www.linguistik-online.org/56_12/
Keating, Elizabeth (2016) The role of the body and space in digital multimodality. In: Alexandra Georgakopoulou and Tereza Spilioti (eds.). The Routledge Handbook of Language and Digital Communication. London: Routledge, 259-272.

Kendon, Adam (1990) Conducting Interaction. Patterns of Behavior in Focused Encounters. (Studies in Interactional Sociolinguistics 7). Cambridge: Cambridge University Press.
Kesselheim, Wolfgang (2017) Die Museumsausstellung – ein Text? Germanistik in der Schweiz – Zeitschrift der Schweizerischen Akademischen Gesellschaft für Germanistik 14, 1-29.
Levinson, Stephen C., and David Wilkins (eds.) (2006) Grammars of Space. Explorations in Cognitive Diversity. Cambridge: Cambridge University Press.
Locher, Miriam A., Andreas H. Jucker and Manuel Berger (2015) Negotiation of space in Second Life newbie interaction. Discourse, Context and Media 9, 34-45.
Löw, Martina (2016) The Sociology of Space. Materiality, Social Structures, and Action. Translated from the German by Donald Goodwin. London: Palgrave Macmillan.
Luhmann, Niklas (1984) Soziale Systeme. Grundriss einer allgemeinen Theorie. Frankfurt am Main: Suhrkamp.
Luhmann, Niklas (2005) Einfache Sozialsysteme. In: Niklas Luhmann (ed.). Soziologische Aufklärung 2. Aufsätze zur Theorie der Gesellschaft. Wiesbaden: VS Verlag für Sozialwissenschaften, 25-47.
Mondada, Lorenza (2009) Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space. Journal of Pragmatics 41.10, 1977-1997.
Ong, Walter H. (1982) Orality and Literacy: The Technologizing of the Word. London: Methuen.
Pentzold, Christian, Claudia Fraas and Stefan Meier (2013) Online-mediale Texte: Kommunikationsformen, Affordanzen, Interfaces. Zeitschrift für germanistische Linguistik 41(1), 81-101.
Piirainen-Marsh, Arja (2012) Organising participation in video gaming activities. In: Ruth Ayaß and Cornelia Gerhardt (eds.). The Appropriation of Media in Everyday Life. (Pragmatics & Beyond New Series 224). Amsterdam: John Benjamins, 197-229.
Rosenbaun, Laura, Sheizaf Rafaeli and Dennis Kurzon (2016a) Participation frameworks in multiparty video chats: Cross-modal exchanges in Public Google Hangouts. Journal of Pragmatics 94, 29–46. DOI: 10.1016/j.pragma.2016.01.003.
Rosenbaun, Laura, Sheizaf Rafaeli and Dennis Kurzon (2016b) Blurring the boundaries between domestic and digital spheres: Competing engagements in Public Google Hangouts. Pragmatics 26, 291–314. DOI: 10.1075/prag.26.2.05ros.
Sacks, Harvey (1984) On doing “being ordinary”. In: J. Maxwell Atkinson and John Heritage (eds.). Structures of Social Action. Studies in Conversation Analysis. Cambridge: Cambridge University Press, 413-429.
Schilt, Kristen, and Laurel Westbrook (2009) Doing gender, doing heteronormativity: “Gender normals,” transgender people, and the social maintenance of heterosexuality. Gender & Society 23(4), 440-464.
Sindoni, Maria Grazia (2012) Mode-switching: How oral and written modes alternate in videochats. In: Mariavita Cambria, Cristina Arizzi and Francesca Coccetta (eds.). Web Genres and Web Tools. Como: Ibis, 141-153.
Sindoni, Maria Grazia (2013) Spoken and Written Discourse in Online Interactions: A Multimodal Approach. New York: Routledge.
Scollon, Ron, and Suzie Wong Scollon (2003) Discourses in Place. Language in the Material World. London: Routledge.
Twitch (2017) Don’t Just Watch, Join In. Available at: https://www.twitch.tv/p/about/.

West, Candace, and Don H. Zimmerman (1987) Doing gender. Gender and Society 1(2), 125-151.
West, Candace, and Don H. Zimmerman (2009) Accounting for doing gender. Gender & Society 23(1), 112-122.